

ОБЩАЯ ПСИХОЛОГИЯ, ПСИХОЛОГИЯ ЛИЧНОСТИ, ИСТОРИЯ ПСИХОЛОГИИ

УДК 159.9.07

Я. А. Ледовая, Р. В. Тихонов, О. Н. Боголюбова

СОЦИАЛЬНЫЕ СЕТИ КАК НОВАЯ СРЕДА ДЛЯ МЕЖДИСЦИПЛИНАРНЫХ ИССЛЕДОВАНИЙ ПОВЕДЕНИЯ ЧЕЛОВЕКА*

Статья содержит описание нового междисциплинарного подхода к сбору индивидуальных психологических, поведенческих и языковых данных в социальных сетях. В описываемой методологии личные данные пользователей социальных сетей (так называемые «цифровые следы», “digital footprints”) собираются с помощью специальных программ и онлайн-приложений. Как правило, участники также заполняют психологические опросники, встроенные в такие приложения. Психологические переменные могут сопоставляться с указанной доступной информацией о поведении пользователей в социальной сети. Все эти данные, получаемые на многотысячных выборках, могут не только анализироваться с помощью классических статистических методов, но и использоваться для построения предсказательных моделей с помощью алгоритмов машинного обучения. Таким образом, психологические переменные (например, личностные особенности, уровень субъективного благополучия и др.) и социально-демографические характеристики могут предсказываться только на основе открытых данных пользователей социальных сетей — текстов, подписок на сообщества и т. п., что является совершенно новым способом получения информации о респондентах. В таких исследовани-

Ледовая Янина Александровна — Санкт-Петербургский государственный университет, Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9; y.ledovaya@spbu.ru

Тихонов Роман Вадимович — аспирант, Санкт-Петербургский государственный университет, Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9; roman.tikhonov@me.com

Боголюбова Ольга Николаевна — кандидат психологических наук, Университет Кларксон, США, 13699, штат Нью-Йорк, Потсдам, а/я 5825; obogolyu@clarkson.edu

Ledovaya Yanina A. — St. Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation; y.ledovaya@spbu.ru

Tikhonov Roman V. — Postgraduate, St. Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation; roman.tikhonov@me.com

Bogolyubova Olga N. — PhD, Clarkson University, P.O.Box 5825, Potsdam, NY, 13699 USA; obogolyu@clarkson.edu

* Работа выполнена при поддержке гранта Санкт-Петербургского государственного университета, проект 8.38.351.2015 «Стресс, здоровье и психологическое благополучие в социальных сетях: кросскультурное исследование».

© Санкт-Петербургский государственный университет, 2017

ях, как правило, участвуют психологи, веб-программисты, а также специалисты по компьютерной лингвистике, анализу данных и машинному обучению. Обсуждаются преимущества и ограничения этой методологии, описаны конкретные подходы к сбору и обработке данных. Представлены некоторые результаты работы пионеров этого направления исследований — участников британского (“Mypersonality.org”) и американского (“World Well-Being Project”) проектов, наиболее масштабно использующих рассматриваемый подход. Библиогр. 47 назв. Табл. 2.

Ключевые слова: социальные сети, Фейсбук, интернет-исследования, сбор данных, цифровые следы, личностные особенности, предсказательные модели, компьютерная лингвистика, междисциплинарный подход.

Y. A. Ledovaya, R. V. Tikhonov, O. N. Bogolyubova

SOCIAL NETWORKS AS A NEW ENVIRONMENT FOR INTERDISCIPLINARY STUDIES OF HUMAN BEHAVIOR

The paper describes a new approach to collecting individual psychological, behavioral and language data from online social networks. Within this approach, personal data (“digital footprints”) are collected by means of special programs and web-applications that are embedded in social networks interfaces or otherwise connected with them. Usually, users provide additional information by answering questions of online surveys embedded in such applications. Psychological variables can be then associated with online behavioral data and other available information. The data of thousands of users can be not only analyzed with traditional statistical methods, but can also be used to build predictive models with machine learning algorithms. Thus, psychological characteristics (personality traits, well-being, etc.) and demographical data can be predicted based on public user information — wall posts, page likes, etc., which is a completely new approach to data collection. Such research projects usually involve multidisciplinary teams of psychologists, web developers, computational linguists and data scientists. Advantages and limitations of this methodology are discussed, as well as the methods of data collection and processing and predictive models building. Key findings of the pioneers of this research direction are presented. These are the findings of the British project “Mypersonality.org” and the USA-based project “World Well-Being Project”. Both are employing the described methodology quite massively. Refs 47. Tables 2.

Keywords: social networks, Facebook, data collection, digital footprints, psychological traits, predictive models, computer linguistics, interdisciplinary approach.

Преимущества и перспективы нового подхода к сбору индивидуальных данных о поведении личности через социальные сети

В 1990-х годах началось повсеместное распространение возможности выхода в Интернет, что повлекло за собой почти неограниченный доступ пользователей к информации и коммуникации друг с другом. Начиная с 2004 г., когда появилась социальная сеть «Фейсбук», социальный и социально-психологический опыт большей части наших современников существенно расширился. Доступ к сети Интернет имеет 51 % жителей Земли [1], от примерно 30 % в Африке до 88 % в Северной Америке, а это в сумме более 3,8 млрд человек. В России уровень проникновения Интернета оценен в 73 % [2]. Не менее двух миллиардов человек ежемесячно пользуются возможностями «Фейсбук» [3], и это число удвоилось за последние пять лет [4]. Помимо «Фейсбук», есть и другие социальные сети, которые получили широкое распространение (табл. 1), однако ни одна из них не достигла масштабов «Фейсбук» по охвату пользователей. Мы сосредоточим свое внимание на дольше всего существующей и доминирующей по охвату и проникновению международной сети «Фейсбук». Большая часть интересующих нас исследований делалась на

основе данных из нее. Кроме того, среди международных социальных сетей именно «Фейсбук» позволяет получить доступ к максимально подробной и персонализированной информации о пользователе через его аккаунт, и именно эти данные наиболее интересны и важны для психологов-исследователей.

Итак, изменился социальный ландшафт, и возникли новые социальные практики (в частности, это расширение круга знакомых и возможность оперативной коммуникации в едином интерфейсе, коллективные обсуждения и простота публичной коммуникации от первого лица с большой аудиторией, распространение информации через друзей и различные виртуальные сообщества, возможность наблюдать за событиями из жизни друзей и знакомых, которыми они делятся, аналогичная возможность делиться значимыми собственными переживаниями в виде текстов, фотографий и видео — список можно продолжать). Ежедневно в социальных сетях люди проводят все больше времени: не менее двух часов в среднем в мире [11] и почти два с половиной часа — в России (данные 2015 г. [12]). Представители молодого поколения в среднем уделяют внимание общению в социальных сетях и мессенджерах еще больше времени, до 160 минут в день [2]. Два часа и более — это большая часть свободного времени, это существенная часть дня. Люди, особенно представители молодого и среднего поколения, действительно вовлечены в эту деятельность очень активно. При этом, по нашим данным, и представители старших поколений активно проводят время в социальных сетях: в выборке из 8396 участников от 18 до 80 лет из России средний возраст составил 45 лет [13; 14].

Таблица 1. Количество активных пользователей социальных сетей в разных странах

Страна	Социальная сеть	Количество активных пользователей (млн человек в месяц)
Россия	«ВКонтакте» [5]	95
	«Одноклассники» [6]	73
Китай	«Qzone» [7]	632
	«Weibo» [8]	313
Международные	«Фейсбук» [4]	2006
	«Инстаграм» [9]	700
	«Твиттер» [10]	328

С начала 2000-х годов представителям социальных наук стало невозможно игнорировать эти перемены в образе жизни современных людей. Некоторые исследователи изучают новые (или заново обнаруживают привычные) формы коммуникации в социальных сетях [15–17], кто-то исследует положительные и отрицательные последствия новых возможностей общения и получения информации, при этом нередко утверждается, что социальные сети приводят к негативным последствиям для психического здоровья, вводится даже алармистское понятие «Фейсбук-аддикции» [18–21].

Но есть категория ученых, которые, сменив парадигму отношения к социальным сетям и рассматривая их не как предмет исследования, а как инструмент, начали использовать социальные сети как платформу для сбора нового типа

данных — следов, которые оставляют пользователи в естественных для себя условиях. Это условия, которые стали для многих привычными и удобными, а значит, поведение в них стало менее контролируемым. Появилось название для таких данных — “digital footprints” — «цифровые отпечатки». Возникли разнообразные автоматизированные формы для их масштабного сбора — это могут быть программы-краулеры [22] (их название произошло от английского слова “crawling” — ползание, они последовательно собирают открытые данные из любых источников в Интернете, в том числе из социальных сетей) или специальные программы-приложения, подключающиеся через программный интерфейс социальной сети (так называемый API — Application Programming Interface) и работающие внутри ее интерфейса, например как автоматические опросники развлекательного характера, возвращающие пользователям обратную связь по результатам тестирования [23]. Исследовательским подразделениям компаний-владельцев социальных сетей доступны вообще все данные, которые оставляют пользователи, включая «закрытые» от посторонних записи (статусы, посты) и фотоальбомы, и даже тексты личных сообщений. Поэтому они могут заниматься еще более сложными, хотя и этически неоднозначными видами анализа таких данных.

Результаты, собранные таким новым способом, лишены ряда методологических недостатков, достаточно известных в психологии и социальных науках, и при этом сохраняют положительные стороны привычных офлайн-исследований [24–26].

Во-первых, на протяжении XX в. страдала внешняя валидность большей части исследований. Их результаты нельзя было легко распространять на всех представителей генеральной совокупности из-за того, что типичными испытуемыми во многих психологических исследованиях были так называемые “WEIRD people” (от “weird” — «странный, неадекватный»): выборки нередко формировались и сейчас формируются по принципу доступности, а доступными для исследовательских проектов университетских ученых из развитых стран западного мира были студенты — представители White, Educated, Industrialized, Rich, Democratic сообществ [27]. Таким образом, представители примерно 12 % мировой популяции были испытуемыми в 80 % публикуемых научных исследований, результаты которых обычно распространяли на условного «среднестатистического испытуемого», т. е. абсолютно любого человека на планете. Используя же данные из социальных сетей, исследователи автоматически получают доступ к гораздо более широкому кругу испытуемых, и их ограничивает теперь в основном лишь степень распространенности Интернета и наличие финансирования для проведения рекламной кампании, привлекающей потенциальных участников [28].

Во-вторых, имея возможность собирать при помощи специальных программ и затем анализировать «цифровые отпечатки» и реальные продукты деятельности пользователей социальных сетей (тексты постов, фотографии, метки местоположения — геотеги, метки тематических ассоциаций — хештеги, списки сообществ, на новости которых подписаны пользователи, списки «друзей» и т. д.), исследователи получают доступ к менее подверженным эффекту социальной желательности и более экологически валидным данным. Стоит отметить, что получение списка друзей возможно не во всех социальных сетях. Сегодня это достаточно несложно делать для социальных сетей «Твиттер» и «ВКонтакте». В сетях «Фейсбук» и «Инстаграм» список друзей получить затруднительно.

В-третьих, многократно возросли размеры выборок исследований, а также сократилась по времени и удешевилась процедура сбора данных. Так, в собираемых в США через сеть «Фейсбук» данных (когда пользователь заполняет, как правило, внутри интерфейса социальной сети ряд психологических опросников, а с его согласия исследователи загружают при этом публично доступную информацию, например упоминавшиеся уже тексты статусов, возраст и пол, местонахождение и информацию о подписках на публичные страницы — “page likes”) стоимость специально настраиваемой рекламы таких опросов (рекламу можно таргетировать на демографически и географически разные группы в особом разделе сети «Фейсбук» для рекламодателей) колеблется в пересчете на одного респондента от 1,51 до 33 долларов, в среднем — 13,75 доллара [24; 29; 30]. По нашему опыту, сбор данных среди российских пользователей гораздо менее затратен.

В-четвертых, для того чтобы получить данные пользователей из практически любой страны, нет необходимости организовывать поездку и искать возможность контактов с представителями другого языка и другой культуры. Достаточно иметь доступ к версиям опросников на нужном языке и сотрудничать с лингвистом, знающим этот язык, — для последующей обработки языковых данных. Также более реалистичными становятся задачи получения данных от маломобильных групп людей (находящихся чаще дома или в учреждениях).

Наконец, в-пятых, на данных такого масштаба можно строить и проверять предсказательные модели о личностных особенностях и поведении людей, например по текстам публичных постов пользователя или по тематике сообществ, на которые он (она) подписан(а) (это так называемые «дешевые», как правило, доступные для сбора при помощи программы-краулера данные), можно предсказывать его (ее) личностные особенности, если до этого было собрано достаточно «дорогих» данных большого количества других пользователей (т. е. таких данных, в которых есть ответы этих пользователей на вопросы психологических опросников, которые можно было сопоставлять для построения моделей с также собранными краулером или при помощи приложения данными о поведении в социальной сети — о текстах, подписках на публичные страницы и т. п.) [31–33]. Мы также рекомендуем посмотреть видеозапись выступления Михала Косински в России, в центральном офисе компании «Сбербанк». В нем он рассказывает о своих исследованиях и принципе создания и работы предсказательных моделей, которые строятся на основе больших данных¹.

Отметим, что в целом качество данных, полученных с использованием социальных сетей и других интернет-технологий, сопоставимо с офлайн-методами сбора данных, что неоднократно отмечалось разными авторами [34–36].

Описываемый подход к сбору данных имеет и свои ограничения.

Во-первых, невозможность контролировать факторы окружающей среды, в которой происходит заполнение опроса, а также степень вовлеченности и внимательности респондента. Вполне возможно, что в момент заполнения опроса кто-то шумит или подсказывает «правильные» варианты ответов, создавая не самую комфортную атмосферу для заполнения психометрических тестов и опросников.

¹ См.: <http://www.sberbanktv.ru/?video=2050> (дата обращения: 01.08.2017).

Во-вторых, достаточно распространено намеренное искажение пользователями информации о своем возрасте и половой принадлежности: респонденты могут заходить в социальные сети через фальшивый профиль, выдавая себя за другого человека.

В-третьих, проблематичность заполнения слишком длинных опросников при сокращении времени типичной онлайн-сессии [24].

И, в-четвертых, использование данных из социальных сетей поднимает новые вопросы исследовательской этики в онлайн-среде. Какую информацию можно собирать и анализировать без ведома пользователей? Можно ли считать, что участники дали информированное согласие на сбор и обработку их персональных данных, если они не глядя нажали на кнопку «Я согласен», чтобы побыстрее перейти к опросу (см. подробнее: [37])? Как повысить осведомленность пользователей о том, какие данные будут доступны исследователям? К сожалению, процедуры рассмотрения интернет-исследований этическими комитетами все еще не до конца отрегулированы. Например, в кодексе АРА, обновленном с января 2017 года, нет упоминаний сбора данных в социальных сетях. А один из самых цитируемых исследователей, во многом пионер этого направления Михал Косински говорит в своих выступлениях о том, что необходимо с осторожностью относиться к открывшимся возможностям работы с большими данными на основе «цифровых следов», потому что предсказания особенностей личности и даже поступков человека, которые становятся все более точными, могут лишить нас приватности и в какой-то мере даже личной безопасности².

Многие из этих ограничений преодолимы за счет качественной предобработки первичных данных: включение в опросник вопросов, оценивающих социальную желательность респондента, использование «вопросов-ловушек», направленных на отсеивание невнимательных участников, контроль времени, выделяемого на вопрос, и исключение слишком быстрых ответов, отслеживание длинных цепочек одинаковых и противоречивых ответов на сходные по смыслу вопросы.

Необходимо отметить, что у рассматриваемого подхода есть не только научные перспективы, связанные с уточнением теоретических моделей личностных черт при помощи новых поведенческих данных или с созданием новых теорий, описывающих личность через ее язык. С опорой на поведение людей в социальных сетях и, в частности, на открытые тексты пользователей возможны варианты практического применения подхода: например, пользователям, в чьих текстах при сканировании программами-краулерами будут обнаруживаться лексемы, темы (кластеры семантически близких слов), эмоциональные валентности, которые до этого были связаны в исследованиях с потенциальным снижением уровня психологического благополучия и угрозой депрессии, суицидальными наклонностями и т. п., можно автоматически (при договоренности с администрацией социальной сети) предъявлять в ленте новостей рекламу служб психологической помощи или объявления, поддерживающие их и разъясняющие природу их плохого самочувствия. С такой практикой столкнулись авторы статьи и их коллеги, когда анализировали случайные изображения из сети «Инстаграм», снабженные подписями-хештегами (так называются метки тематических ассоциаций, которые

² См.: <http://www.sberbanktv.ru/?video=2050> (дата обращения: 01.08.2017).

произвольно ставят сами пользователи). Анализировались и кодировались изображения, сопровождавшиеся хештегами, описывающими негативные эмоции и переживания, — #depression, #anxiety, #fear, #stress, #worry (депрессия, тревога, страх, стресс, беспокойство) — на английском и русском языках соответственно. До половины изображений, сопровождавшихся меткой “#depression” вместо собственно изображений, фотографий, содержали «мотиваторы» — картинки со словами поддержки, объяснением причин и симптомов депрессии и номерами телефонов или адресами сайтов организаций, оказывающих психологическую помощь (зачастую бесплатную). При этом изображения, сопровождавшиеся аналогичным хештегом на русском языке, не содержали подобных текстов, которые могли бы выполнять функции психологической помощи. Анализируя этот факт, мы предположили, что культура оказания психологической помощи тем, кто может сообщать через свои публикации в социальной сети об ухудшении психического здоровья или искать именно в социальной сети поддержку, еще недостаточно хорошо развита в русскоязычном мире по сравнению с англоязычным [38].

Возможность доступа к записям пользователей за все время, начиная с момента регистрации в социальной сети, позволяет исследователям из области общественного здоровья ретроспективно проводить лонгитюдный анализ и восстанавливать особенности образа жизни пользователей, делать уточнения для постановки диагнозов [26]. Говоря о прикладном применении данных из социальных сетей, можно также упомянуть исследование, в котором до 70 % пациентов, ожидающих медицинской помощи в приемном покое больницы, имеющих аккаунт в социальной сети и согласившихся участвовать в устном опросе, дают разрешение пользоваться автоматически подключаемыми к их медицинским картам данными из их социально-сетевых аккаунтов — для уточнения диагноза и проведения научных исследований [39].

**“MyPersonality.org” (Кембриджский университет),
“World Well-Being Project” (Университет Пенсильвании)
и исследования на основе данных пользователей социальных сетей
«Фейсбук» и «Твиттер»**

Перейдем к тому, как именно исследователи собирают одновременно психологические и языковые данные, а также данные о поведении в социальной сети, напрямую обращаясь к пользователям социальных сетей, как анализируют их и какие результаты получают. Насколько нам известно, сегодня в мире существует лишь один крупный научный проект, в котором психологи, лингвисты, программисты и специалисты по обработке данных используют возможности сбора информации в социальных сетях для оценки психологического благополучия и физического здоровья с опорой на анализ языка социальных медиа. Это проект Центра позитивной психологии Университета Пенсильвании “World Well-Being Project” [40], стартовавший в 2011 г. [25]. В нем используются данные из уникальной базы пользователей социальной сети «Фейсбук», собранные благодаря идее и усилиям научных сотрудников в тот момент Кембриджского университета Дэвида Стилвелла и Михала Косински (сейчас Михал Косински работает в Стэнфордском университете). Они создали программу,

которая работала на их сайте с 2007 по 2012 г. и позволяла различным (англоязычным) людям участвовать в заполнении любого количества из примерно двух десятков распространенных психологических опросников, возвращавших им обратную связь с результатами. Кроме того, с согласия этих пользователей программа загружала демографическую информацию из их профиля и публично доступные тексты их статусов из сети «Фейсбук», а также «лайки» страниц. У проекта есть сайт (URL: <http://mypersonality.org/wiki/doku.php>), на котором некоторые данные доступны для анализа зарегистрированным сторонним исследователям. Более 7,5 млн уникальных пользователей «Фейсбук» участвовали в тестировании на сайте и оставили данные своих профилей. Стилвелл и Косински часто являются соавторами работ упомянутой лаборатории Университета Пенсильвании, которой руководит Мартин Селигман. Всего в команде проекта “World Well-Being Project” около 20 постоянных участников и около 15 коллабораторов со всего мира. Работают они во многом благодаря спонсорской помощи фонда “Templeton Religion Trust” (на значимость проекта указывает то, что с 2013 г. этот фонд поддержал проекты Мартина Селигмана на сумму, превышающую 9 млн долларов [41]). Они работают с данными, собираемыми также и в социальной сети «Твиттер». Эти данные, в отличие от данных, собираемых с помощью приложений для сети «Фейсбук», гораздо более доступны: руководство сети «Твиттер» открыто к коллаборациям и дает возможность скачивать их через API и даже продает массивы данных для исследований [42]. Но социально-демографической информации в профилях этой сети гораздо меньше, тексты существенно короче, в ней много ботов (искусственных аккаунтов, которые не ведут люди), и работа с данными «Твиттер» скорее будет представлять интерес для лингвистов и специалистов, отслеживающих каналы распространения информации и связи между участниками сети, — социологов, политологов (исследовать социальные связи пользователей, например, получив список их «друзей», в настоящее время затруднительно через API «Фейсбук»). Надо отметить, что на сайте пенсильванского проекта (URL: <http://wwbp.org/>) и на личном сайте Михала Косински (URL: <http://www.michalkosinski.com/home>) выложены в открытом доступе десятки публикаций, сделанных на основе работы авторов с данными из социальных сетей.

Основные результаты проекта “World Well-Being Project” связаны с тем, какие языковые проявления (слова, эмоциональная валентность) свойственны людям с различными психологическими чертами или демографическими характеристиками. Находя сперва корреляции между психологическими, демографическими и языковыми данными, а также особенностями поведения в социальной сети (фотографиями профиля, степенью активности и т. д.), исследователи затем строят предсказательные модели. Приведем несколько примеров. В одной из публикаций описывается, как по смысловым характеристикам публикуемых в сети «Твиттер» постов и фотографий профиля пользователя предсказать нарциссизм и психопатию: по полученным данным, люди с более высоким нарциссизмом публикуют позитивную информацию, а на фотографии профиля будут, скорее всего, они сами, причем улыбаясь, в то время как люди с более высоким уровнем психопатии используют в текстах широкий спектр негативных эмоций, связанных с сообщениями о разных видах насилия [43].

На основании анализа текстов статусов (так в сети «Фейсбук» называют сообщения, публикуемые пользователями на своих «стенах» для друзей или всеобщего доступа) исследователи из этого коллектива смогли, в частности, описать отличия «женского» языка от «мужского» по нескольким параметрам: женщины (если судить по полу, приписанному самими пользователями в аккаунтах) чаще пишут о семье, друзьях и различных событиях социальной жизни (анализировались тематические кластеры семантически связанных слов (“topics”), используемых в статусах), а мужчины в текстах ругаются, выражают гнев, участвуют в дискуссиях о политике, спорте, музыке и видеоиграх и реже рассуждают о людях [44].

В статье, описывающей возможности предсказывать уровень благополучия (удовлетворенности жизнью) с опорой на язык социальных сетей, анализировались данные более 2000 человек, добровольно участвовавших в их сборе через сайт Д. Стилвелла и М. Косински, заполнивших опросник удовлетворенности жизнью Э. Динера [45] и давших доступ к параметрам своего аккаунта (в том числе к текстам статусов в сети «Фейсбук» [46]). Исследователи, традиционно для такого рода работ, разделили выборку, 80 % данных которой участвовали в создании и обучении нескольких компьютерно-лингвистических моделей, предсказывавших уровень удовлетворенности жизнью с опорой на тексты, а 20 % были тестовыми, на которых эти модели проверялись («метод кросс-валидации с пятью разбиениями»). Авторам удалось достичь эффективности предсказания уровня удовлетворенности жизнью на отметке 0,566 (в модель вошли *n*-граммы, тематические кластеры слов (“topics”) и отдельные лексемы). Более всего коррелировали с результатами опросника SWLS тематические кластеры. Среди наиболее коррелирующих были четыре положительно с ним связанных («Эмоциональная вовлеченность» — «потрясающе», «супер», «завтра»; «Социально-гражданская вовлеченность» — «встреча», «конференция», «персонал», «посещать»; «Профессиональные ценности» — «управление», «навыки», «учеба», «бизнес»; «Теплые отношения» — «семья», «друзья», «благодарен», «с любовью») и два отрицательно связанных («Отчуждение» как антипод «Вовлеченности» — «скучно», «скука», «текст» и кластер, состоящий из обценных слов и выражений — яркий пример предиктора низкого уровня удовлетворенности жизнью — предиктор со знаком минус)³. Исследователи сопоставляют обобщенные и осмысленные значения этих кластеров с имеющимися в психологической литературе коррелятами удовлетворенности жизнью, традиционно определяемыми через психологические опросники или наблюдаемое поведение людей, и приходят к выводу о том, что сочетание лингвистических и психологических данных может стать не только материалом для сугубо прикладной задачи предсказывать одно через другое (традиционной для компьютерной лингвистики), но и для более глубокого теоретического осмысления самого понятия «удовлетворенность жизнью», его причин, форм проявления и последствий.

Самая цитируемая работа Михала Косински и его соавторов (Дэвида Стилвелла из Кембриджского университета и сотрудника компании «Майкрософт» Тора Грейпела), упомянутая с апреля 2013 г. уже в 226 статьях, называется “Private traits and attributes are predictable from digital records of human behavior” («Инди-

³ Примеры слов и названия кластеров даны здесь в переводе авторов статьи.

видуальные черты и характеристики можно предсказывать на основе цифровых данных о поведении человека») [47]. В ней был представлен по-своему революционный подход к работе с данными о поведении людей в сети Интернет, в частности в социальной сети «Фейсбук». Желавшие пройти психологическое тестирование с обратной связью на сайте проекта “MyPersonality.org” в 40 % случаев давали разрешение исследователям получить ряд характеристик их аккаунта в сети «Фейсбук» (пол и возраст, семейное положение и сексуальная ориентация (в профиле сети «Фейсбук» можно обозначить, что пользователь «заинтересован встречаться с... (мужчинами или женщинами)»), политические и религиозные взгляды, список публичных страниц (сообществ), на которые был подписан пользователь (это и были пресловутые “Facebook likes”, исходя из которых можно с достаточно высокой вероятностью предсказывать скрытые психологические и демографические характеристики пользователей), а также на тот момент было возможно получить список друзей). Кроме того, несколько вопросов задавалось им дополнительно (о курении, употреблении алкоголя и наркотиков и о том, развелись ли их родители до достижения ими 21 года).

Из 25 доступных психологических опросников для рассматриваемой статьи были взяты результаты трех — одной из версий «Большой пятерки личностных свойств», теста Дж.Равена и опросника удовлетворенности жизнью Э.Динера SWLS. Данные разного количества респондентов, от 766 (вопрос о разводе родителей) до 54 373 (оценки по тесту «Большой пятерки»), попали в выборку. 17 601 респондент дал доступ к профилю в сети «Фейсбук». По нескольким тысячам случайно выбранных фотографий из профилей, просмотренных «вручную», делался вывод об этнической принадлежности респондента (европейской или африканской). Среднее количество «лайков» в выборке равнялось 170 (еще раз подчеркнем, что этот параметр отражает устойчивые интересы пользователей, от этих публичных страниц (сообществ) они регулярно получают новости, и это не разовое выражение симпатии к новой фотографии друга или обновлению его статуса).

Далее все множество пользователей и все возможные публичные страницы («лайки») были помещены в матрицу, размерность которой была многократно уменьшена при помощи процедуры сингулярного разложения (“singular-value decomposition, SVD”): можно сказать, что почти 56 000 разных публичных страниц, на которые были подписаны 17 600 респондентов, были сведены к 100 «компонентам» — страницам («лайкнутым» страницам), наиболее часто встречающимся в выборке и наиболее точно в итоге предсказывающим важные психологические и демографические параметры. Затем для предсказания количественных переменных (таких, как интеллект или удовлетворенность жизнью) использовалась линейная регрессионная модель, а для бинарных (пол, сексуальная ориентация) — логистическая регрессия. Модели были обучены достаточно типичным в машинном обучении методом кросс-валидации с десятью разбиениями (10-fold cross-validation). Он позволяет оценить поведение аналитической модели на независимых данных. Предсказанные величины параметров затем сопоставлялись с реальными.

Точность предсказания бинарных переменных с опорой на «лайки» страниц была порою очень высокой. Так, вероятность точного определения расы — 0,95, пола — 0,93, политических предпочтений (демократических или республиканских) — 0,85, вероисповедания (христианство или ислам) — 0,82. Гомосексуаль-

ность мужчин — 0,88, женщин — 0,75. От 0,65 до 0,7 — вероятность точного определения семейного статуса, пристрастий к курению, алкоголю, наркотикам. Развод родителей до достижения совершеннолетия определялся хуже всего — с вероятностью 0,6, но здесь было меньше всего респондентов и их данных.

Точность предсказания количественных параметров (выраженная как коэффициент корреляции между предсказанным и реальным параметрами) была в целом ниже, чем бинарных. Возраст был предсказан точнее всего — с вероятностью 0,75. Приблизительно между 0,4 и 0,5 получились вероятности предсказания интеллекта, экстраверсии, открытости новому опыту, количества друзей. С коэффициентами от 0,2 до 0,3 коррелировали реальные и предсказанные эмоциональная стабильность, доброжелательность и сознательность. А предсказать уровень удовлетворенности жизнью по шкале Э. Динера при помощи моделей, построенных на данных о подписках на страницы сообществ, можно сказать, не удалось: точность предсказания удовлетворенности жизнью составила 0,17. Исследователи рассуждают, что причиной может быть тот факт, что «лайки» страниц отражают слишком устойчивые интересы, а в шкалу удовлетворенности жизнью заложены в том числе вопросы о достаточно нестабильных сменах настроения. Ниже представлен список публичных страниц, подписки на которые теснее всего связаны с тем или иным полюсом бинарных шкал, а также с высокими и низкими значениями по количественным шкалам. Поскольку все участники исследования — жители США, этот список культурно специфичен. Приведем несколько примеров (табл. 2).

Таблица 2. Примеры некоторых публичных страниц, наиболее точно предсказывающих психологические, демографические и поведенческие характеристики пользователей сети «Фейсбук» (данные статьи М. Косински и соавторов) [47]

Параметр	Уровень	Некоторые наиболее точно предсказывающие параметр «лайки» страниц	Точность
Интеллект	высокий	Mozart Science (журнал) Curly Fries	0,39
	низкий	Sephora I Love Being A Mom Harley Davidson	
Удовлетворенность жизнью	высокая	Proud To Be Christian Swimming Being Conservative	0,17
	низкая	Gorillaz Science Ipod	
Открытость опыту	высокая	Oscar Wilde Charles Bukowski Leonard Cohen	0,43
	низкая	I Don't Read The Bachelor Oklahoma State University	

Параметр	Уровень	Некоторые наиболее точно предсказывающие параметр «лайки» страниц	Точность
Эмоциональная стабильность	высокая	Getting Money Mountain Biking Physics / Engineering	0,3
	низкая	Sometimes I Hate Myself Emo Kurt Donald Cobain	
Пол	женский	TV Fantastic Gillette Venus Proud To be A Mom	0,93
	мужской	Sportscenter Starcraft Bruce Lee	
Возраст	пожилой	Cup Of Joe For A Joe Proud To Be A Mom Fly The American Flag	0,75
	молодой	Walt Disney Records I Hate My ID Photo 293 Things To Do In Class When You Are Bored	
Этническое происхождение	африканское	I Support My President Next Friday Love And Basketball	0.95
	европейское	I Come From A Town Where A Traffic Jam Is 4 Cars Behind A Tractor Harley Davidson David Bowie Official	
Отношение к алкоголю	употребляют	Dear Liver Thank You're A Champ Trying To Figure Out If Its A Cop Car Belvedere Vodka	0,7
	не употребляют	Honda How To Make A Girl Smile <3<3 When Your Fortune Cookie Knows What's Up	

Коротко прокомментируем результаты двух последних исследований. Мы видим, что такие показатели, как кластеры слов и «лайки» публичных страниц, полученные силами специалистов по компьютерной лингвистике и анализу данных, могут дать психологам немало новых идей о том, как в реальной жизни, в собственной речи или через призму стабильного интереса к социально-сетевым сообществам люди переживают, осмысливают и «озвучивают» собственные психологические черты. Такими примерами можно по-новому наполнить содержание понятий, привычных для психологов и нередко описываемых в достаточно абстрактных категориях.

Также можно отметить, что для предсказания уровня удовлетворенности жизнью, по данным этих двух исследований, более эффективными предикторами являются тексты пользователей, а не их подписки на публичные страницы. Можно предположить, что удовлетворенность жизнью — достаточно интимно переживаемая сущность, тесно связанная с «я», поэтому прямая речь субъекта может быть более точным предиктором, чем выражающие скорее социально обусловленные интересы личности подписки на публичные страницы в социальной сети.

Заключение

Распространение социальных сетей в мире нарастает, а это значит, что все большее количество людей будет доступно для участия в исследованиях через социальные сети. Мы обозначили изменения в социально-коммуникативных практиках и образе жизни большей части современных людей, появившиеся в связи с распространением социальных сетей. Также мы затронули вопросы перспектив и ограничений существующих методологических и методических подходов к исследованиям «цифровых следов», которые оставляют пользователи социальных сетей. Практически все исследования такого рода связаны с использованием возможностей сетей «Фейсбук» и «Твиттер». При этом в России, судя по результатам поиска в научной базе РИНЦ, практически не проводятся такого рода психологические исследования с использованием API социальных сетей и возможностями больших выборок.

Технологии машинного обучения, стремительно развивающиеся последние 2–4 года, дают новые и очень масштабные возможности. Например, они позволяют предсказывать (с определенной вероятностью) на основе получаемых массивов поведенческих данных наличие определенных черт личности исходя из текстов пользователей и их подписок на публичные страницы. Это новый для психологии подход, и он дает возможности как для прикладного применения таких моделей, так и для фундаментальных теоретических инсайтов относительно наличия устойчивых связей между скрытыми психологическими чертами и внешне наблюдаемыми продуктами деятельности в публичном социально-сетевом пространстве.

Использование различных аспектов языковых (лингвистических) коррелятов психологических особенностей людей (от частоты написания постов и богатства словаря до специфических лексем, эмоциональной валентности и тематических кластеров слов) открывает интересные перспективы для прикладных и фундаментальных междисциплинарных работ, в которых психологи могут ставить задачи, а компьютерные лингвисты и специалисты по работе с данными и машинному обучению — заниматься разработкой алгоритмов и метрик для их решения.

Мы хотим еще раз подчеркнуть важность междисциплинарного сотрудничества в рамках такого рода исследований: в научных проектах, собирающих и анализирующих индивидуальные данные пользователей социальных сетей, необходимы web-программисты для работы над приложением, лингвисты для многоступенчатой и разнонаправленной обработки текстовых данных, data-аналитики и специалисты по машинному обучению. Безусловно, важна роль психологов. Именно они ставят задачи и ориентируются на свои знания о поведении человека. Но междисциплинарное сотрудничество означает, что психологам необходимо разви-

вать (хотя бы на начальном уровне) компетенции в области лингвистики и компьютерной лингвистики, подходов к работе с большими массивами данных, основ и логики машинного обучения. Тогда работа таких команд обещает чрезвычайно продуктивное сотрудничество.

Литература

1. Miniwatts Marketing Group. World Internet usage and population statistics: June, 30 2017 // Internet World Stats. 2017. URL: <http://www.internetworldstats.com/stats.htm> (дата обращения: 01.08.2017).
2. Mander J., McGrath F. GWI Social Summary Q1 2017 // GlobalWebIndex. 2017. URL: <https://www.globalwebindex.net> (дата обращения: 10.08.2017).
3. Facebook Inc. Company Info // Facebook Newsroom. URL: <https://newsroom.fb.com/company-info/> (дата обращения: 10.08.2017).
4. Statista Inc. Facebook users worldwide 2008–2017. 2017. URL: <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/> (дата обращения: 10.08.2017).
5. ВКонтатке. О компании. URL: <https://vk.com/about> (дата обращения: 12.08.2017).
6. Интерфакс-Украина. «Одноклассники» в 2016 году увеличили количество пользователей на 10 %. 2017. URL: <http://interfax.com.ua/news/economic/396363.html> (дата обращения: 05.08.2017).
7. DMR. 5 amazing Qzone stats and facts (February 2017). URL: <http://expandedramblings.com/index.php/business-directory/19888/qzone/> (дата обращения: 10.08.2017).
8. DRM. 61 amazing Weibo statistics and facts (March 2017). URL: <http://expandedramblings.com/index.php/weibo-user-statistics/> (дата обращения: 10.08.2017).
9. Statista Inc. Instagram: number of monthly active users 2013–2017. 2017. URL: <https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/> (дата обращения: 10.08.2017).
10. Statista Inc. Twitter: number of monthly active users 2010–2017. 2017. URL: <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/> (дата обращения: 10.08.2017).
11. Mander J. Daily time spent on social networks rises to over 2 hours // GlobalWebIndex. 2017. URL: <http://blog.globalwebindex.net/chart-of-the-day/daily-time-spent-on-social-networks/> (дата обращения: 12.08.2017).
12. Исследовательский холдинг Ромир. Социально-сетевая жизнь. 2015. URL: http://romir.ru/studies/670_1432155600/ (дата обращения: 21.08.2017).
13. Bogolyubova O., Tikhonov R., Ivanov V., Panicheva P., Ledovaya Y. Violence exposure, posttraumatic stress, and subjective well-being in a sample of Russian adults // Journal of Interpersonal Violence. 2017. URL: <https://doi.org/10.1177/0886260517698279> (дата обращения: 21.08.2017).
14. Panicheva P., Ledovaya Y., Bogolyubova O. Lexical, Morphological and semantic correlates of the Dark Triad personality traits in Russian Facebook texts // Conference Paper. AINL FRUCT 2016. Saint-Petersburg, Russia. 2016. URL: <http://ainlconf.ru/2016/materials> (дата обращения: 21.08.2017).
15. Gosling S. D., Mason W. Internet Research in Psychology // Annual Review of Psychology. 2015. Vol. 66. P. 877–902.
16. Dunbar R. I. M., Arnaboldi V., Conti M., Passarella A. The structure of online social networks mirrors those in the offline world // Social Networks. 2015. Vol. 43. P. 39–47.
17. Tifentale A., Manovich L. Selficity: Exploring Photography and Self-Fashioning in Social Media // Postdigital Aesthetics. London: Palgrave Macmillan UK, 2015. P. 109–122.
18. Gonzales A. L., Hancock J. T. Mirror, Mirror on my facebook wall: effects of exposure to Facebook on self-esteem // Cyberpsychol Behav Soc Netw. 2011. Vol. 14, N 1–2. P. 79–83.
19. Kim J., Lee J.-E. R. The Facebook Paths to Happiness: Effects of the Number of Facebook Friends and Self-Presentation on Subjective Well-Being // Cyberpsychol Behav Soc Netw. 2011. Vol. 14, N 6. P. 359–364.
20. Kross E. et al. Facebook Use Predicts Declines in Subjective Well-Being in Young Adults // PLoS One. 2013. Vol. 8, N 8. P. e69841.
21. Ryan T., Chester A., Reece J., Xenos S. The uses and abuses of Facebook: A review of Facebook addiction // Journal of Behavioral Addictions. 2014. Vol. 3, N 3. P. 133–148.
22. Butakov N., Petrov M., Radice A. Multitenant Approach to Crawling of Online Social Networks // Procedia Computer Science. 2016. Vol. 101. P. 115–124.
23. Ледовая Я. А., Тихонов Р. В., Боголюбова О. Н., Казенная Е. В., Сорокина Ю. Л. Отчужде-

ние моральной ответственности: психологический конструкт и методы его измерения // Вестник С.-Петерб. ун-та. Серия 16. Психология и педагогика. 2016. Т. 16, № 4. С. 23–39.

24. *Kosinski M. et al.* Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines // *Am Psychology*. 2015. Vol. 70, N 6. P. 543–556.

25. *Kern M. L. et al.* Gaining insights from social media language: Methodologies and challenges // *Psychology Methods*. 2016. Vol. 21, N 4. P. 507–525.

26. *Inkster B., Stillwell D., Kosinski M., Jones P. A.* A decade into Facebook: where is psychiatry in the digital age? // *The Lancet Psychiatry*. 2016. Vol. 3, N 11. P. 1087–1090.

27. *Azar B.* Are your findings “WEIRD”? // *Monitor on Psychology*. 2010. Vol. 41, N 5. P. 11.

28. *Gosling S. D., Sandy C. J., John O. P., Potter J.* Wired but not WEIRD: The promise of the Internet in reaching more diverse samples // *Behavioral and Brain Sciences*. 2010. Vol. 33, N 2–3. P. 94–95.

29. *Batterham P. J.* Recruitment of mental health survey participants using Internet advertising: content, characteristics and cost effectiveness // *International journal of methods in psychiatric research*. 2014. Vol. 23, N 2. P. 184–191.

30. *Richiardi L., Pivetta E., Merletti F.* Recruiting Study Participants Through Facebook // *Epidemiology*. 2012. Vol. 23, N 1. P. 175.

31. *Schwartz H. A., Ungar L. H.* Data-Driven Content Analysis of Social Media // *The ANNALS of the American Academy of Political and Social Science*. 2015. Vol. 659, N 1. P. 78–94.

32. *Panicheva P., Mirzagitova A., Ledovaya Y.* Semantic Feature Aggregation for Gender Identification in Russian Facebook // *Proceedings of the AINL*. 2017. (In press).

33. *Moskvichev A., Menshov S., Dubova M., Filchenkov A.* Using Linguistic Activity In Social Networks To Predict and Interpret Dark Psychological Traits // *Proceedings of the AINL*. 2017. (In press)

34. *Casler K., Bickel L., Hackett E.* Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing // *Computers in Human Behavior*. 2013. Vol. 29, N 6. P. 2156–2160.

35. *Ramsey S. R., Thompson K. L., McKenzie M., Rosenbaum A.* Psychological research in the internet age: The quality of web-based data // *Computers in Human Behavior*. 2016. Vol. 58. P. 354–360.

36. Одайник А. С., Четвериков А. А. Проведение экспериментальных психологических исследований в сети Интернет // Психология XXI века: Мат-лы Междунар. науч.-практ. конф. молодых ученых «Психология XXI века». 21–23 апреля 2011 г., Санкт-Петербург / под ред. О. Ю. Щелковой. СПб.: Изд-во СПбГУ, 2011. С. 85–87.

37. British Psychological Society. Ethics Guidelines for Internet-mediated Research. 2017. URL: <http://www.bps.org.uk/publications/policy-and-guidelines/research-guidelines-policy-documents/research-guidelines-poli> (дата обращения: 20.07.2017).

38. Боголюбова О. Н., Ледовая. Я. А., Чурилова А. Г. Репрезентации психологического дистресса в сети «Инстаграм» // Ананьевские чтения — 2016. Психология: вчера, сегодня, завтра: Мат-лы междунар. науч. конф. 25–29 октября 2016 г.: В 2 т. / под ред. А. В. Шаболаса и др. СПб.: Айсинг, 2016. Т. 2. С. 125–126.

39. *Padrez K. A. et al.* Linking social media and medical record data: a study of adults presenting to an academic, urban emergency department // *BMJ quality & safety*. 2016. Vol. 25, N 6. P. 414–423.

40. Penn Positive Psychology Center. World well-being project. URL: <http://wwbp.org/about.html> (дата обращения: 22.08.2017).

41. John Templeton Foundation. Grant Database. 2017. URL: <https://www.templeton.org/grants/grant-database> (дата обращения: 22.08.2017).

42. Gnip Inc. Enterprise access to Twitter data. URL: <https://gnip.com/sources/twitter/> (дата обращения: 22.08.2017).

43. *Preotiuc-Pietro D., Carpenter J., Giorgi S., Ungar L.* Studying the Dark Triad of personality through Twitter behavior // *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management — CIKM '16*. New York, New York: ACM Press, 2016. P. 761–770.

44. *Park G. et al.* Women are Warmer but No Less Assertive than Men: Gender and Language on Facebook // *PLOS ONE*. 2016. Vol. 11, N 5. P. e0155885.

45. *Diener E., Emmons R. A., Larsen R. J., Griffin S.* The Satisfaction With Life Scale // *Journal of Personality Assessment*. 1985. Vol. 49, N 1. P. 71–75.

46. *Schwartz H. A. et al.* Predicting individual well-being through the language of social media // *Biocomputing 2016: Proceedings of the Pacific Symposium*. 2016. P. 516–527.

47. *Kosinski M., Stillwell D., Graepel T.* Private traits and attributes are predictable from digital records of human behavior // *Proceedings of the National Academy of Sciences (PNAS)*. 2013. Vol. 110, N 15. P. 5802–5805.

References

1. Miniwatts Marketing Group. *World Internet usage and population statistics: June, 30 2017. Internet World Stats*. 2017. Available at: <http://www.internetworldstats.com/stats.htm> (accessed: 01.08.2017).
2. Mander J., McGrath F. *GWI Social Summary Q1 2017. GlobalWebIndex*. 2017. Available at: <https://www.globalwebindex.net> (accessed: 10.08.2017).
3. Facebook Inc. *Company Info. Facebook Newsroom*. Available at: <https://newsroom.fb.com/company-info/> (accessed: 10.08.2017).
4. Statista Inc. *Facebook users worldwide 2008–2017*. 2017. Available at: <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/> (accessed: 10.08.2017).
5. VKontakte. *O kompanii [VK.com: About us]*. Available at: <https://vk.com/about> (accessed: 12.08.2017). (In Russian)
6. Interfaks-Ukraina. *“Odnoklassniki” v 2016 godu uvelichili kolichestvo pol’zovatelei na 10% [“Odnoklassniki” increased in 2016 the number of users by 10%]*. 2017. Available at: <http://interfax.com.ua/news/economic/396363.html> (accessed: 05.08.2017). (In Russian)
7. DMR. *5 amazing Qzone stats and facts (February 2017)*. Available at: <http://expandedramblings.com/index.php/business-directory/19888/qzone/> (accessed: 10.08.2017).
8. DRM. *61 amazing Weibo statistics and facts (March 2017)*. Available at: <http://expandedramblings.com/index.php/weibo-user-statistics/> (accessed: 10.08.2017).
9. Statista Inc. *Instagram: number of monthly active users 2013–2017*. 2017. Available at: <https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/> (accessed: 10.08.2017).
10. Statista Inc. *Twitter: number of monthly active users 2010–2017*. 2017. Available at: <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/> (accessed: 10.08.2017).
11. Mander J. Daily time spent on social networks rises to over 2 hours. *GlobalWebIndex*. 2017. Available at: <http://blog.globalwebindex.net/chart-of-the-day/daily-time-spent-on-social-networks/> (accessed: 12.08.2017).
12. Research holding Romir. *Sotsial’no-setevaya zhizn’ [Life in social networks]*. 2015. Available at: http://romir.ru/studies/670_1432155600/ (accessed: 21.08.2017). (In Russian)
13. Bogolyubova O., Tikhonov R., Ivanov V., Panicheva P., Ledovaya Y. Violence exposure, posttraumatic stress, and subjective well-being in a sample of Russian adults. *Journal of Interpersonal Violence*, 2017. Available at: <https://doi.org/10.1177/0886260517698279> (accessed: 21.08.2017).
14. Panicheva P., Ledovaya Y., Bogolyubova O. Lexical, Morphological and semantic correlates of the Dark Triad personality traits in Russian Facebook texts. *Conference Paper. AINL FRUCT 2016. Saint-Petersburg, Russia*. 2016. Available at: <http://ainlconf.ru/2016/materials> (accessed: 21.08.2017).
15. Gosling S. D., Mason W. Internet Research in Psychology. *Annual Review of Psychology*, 2015, vol. 66, pp. 877–902.
16. Dunbar R. I. M., Arnaboldi V., Conti M., Passarella A. The structure of online social networks mirrors those in the offline world. *Social Networks*, 2015, vol. 43, pp. 39–47.
17. Tifentale A., Manovich L. Selfcity: Exploring Photography and Self-Fashioning in Social Media. *Postdigital Aesthetics*. London, Palgrave Macmillan UK, 2015, pp. 109–122.
18. Gonzales A. L., Hancock J. T. Mirror, Mirror on my facebook wall: effects of exposure to Facebook on self-esteem. *Cyberpsychol Behav Soc Netw.*, 2011, vol. 14, no. 1–2, pp. 79–83.
19. Kim J., Lee J.-E. R. The Facebook Paths to Happiness: Effects of the Number of Facebook Friends and Self-Presentation on Subjective Well-Being. *Cyberpsychol Behav Soc Netw.*, 2011, vol. 14, no. 6, pp. 359–364.
20. Kross E. et al. Facebook Use Predicts Declines in Subjective Well-Being in Young Adults. *PLoS One*, 2013, vol. 8, no. 8, p. e69841.
21. Ryan T., Chester A., Reece J., Xenos S. The uses and abuses of Facebook: A review of Facebook addiction. *Journal of Behavioral Addictions*, 2014, vol. 3, no. 3, pp. 133–148.
22. Butakov N., Petrov M., Radice A. Multitenant Approach to Crawling of Online Social Networks. *Procedia Computer Science*, 2016, vol. 101, pp. 115–124.
23. Ledovaya Y. A. et al. Otchuzhdenie moral’noi otvetstvennosti: psikhologicheskii konstrukt i metody ego izmereniya [Moral disengagement: the psychological construct and its measurement]. *Vestnik of St. Petersburg University. Series 16*, 2016, vol. 16, no. 4, pp. 23–39. (In Russian)

24. Kosinski M. et al. Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. *Am Psychology*, 2015, vol. 70, no. 6, pp. 543–556.
25. Kern M. L. et al. Gaining insights from social media language: Methodologies and challenges. *Psychology Methods*, 2016, vol. 21, no. 4, pp. 507–525.
26. Inkster B., Stillwell D., Kosinski M., Jones P. A decade into Facebook: where is psychiatry in the digital age? *The Lancet Psychiatry*, 2016, vol. 3, no. 11, pp. 1087–1090.
27. Azar B. Are your findings “WEIRD”? *Monitor on Psychology*, 2010, vol. 41, no. 5, p. 11.
28. Gosling S. D., Sandy C. J., John O. P., Potter J. Wired but not WEIRD: The promise of the Internet in reaching more diverse samples. *Behavioral and Brain Sciences*, 2010, vol. 33, no. 2–3, pp. 94–95.
29. Batterham P. J. Recruitment of mental health survey participants using Internet advertising: content, characteristics and cost effectiveness. *International journal of methods in psychiatric research*, 2014, vol. 23, no. 2, pp. 184–191.
30. Richiardi L., Pivetta E., Merletti F. Recruiting Study Participants Through Facebook. *Epidemiology*, 2012, vol. 23, no. 1, p. 175.
31. Schwartz H. A., Ungar L. H. Data-Driven Content Analysis of Social Media. *The ANNALS of the American Academy of Political and Social Science*, 2015, vol. 659, no. 1, pp. 78–94.
32. Panicheva P., Mirzagitova A., Ledovaya Y. Semantic Feature Aggregation for Gender Identification in Russian Facebook. *Proceedings of the AINL 2017*. (In press).
33. Moskvichev A., Menshov S., Dubova M., Filchenkov A. Using Linguistic Activity In Social Networks To Predict and Interpret Dark Psychological Traits. *Proceedings of the AINL 2017*. (In press).
34. Casler K., Bickel L., Hackett E. Separate but equal? A comparison of participants and data gathered via Amazon’s MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 2013, vol. 29, no. 6, pp. 2156–2160.
35. Ramsey S. R., Thompson K. L., McKenzie M., Rosenbaum A. Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, 2016, vol. 58, pp. 354–360.
36. Odainik A. S., Chetverikov A. A. Provedenie eksperimental’nykh psikhologicheskikh issledovaniy v seti internet [Conducting experimental psychological research in the internet]. *Psikhologiya XXI veka: Materialy Mezhdunarodnoi nauchno-prakticheskoi konferentsii molodykh uchenykh* [Psychology of XXI century: Proceedings of international scientific-practical conference of young researchers]. Ed. by O. Y. Shchelkova. St. Petersburg, St. Petersburg University Press, 2011, pp. 85–87. (In Russian)
37. British Psychological Society. *Ethics Guidelines for Internet-mediated Research*. 2017. Available at: <http://www.bps.org.uk/publications/policy-and-guidelines/research-guidelines-policy-documents/research-guidelines-poli> (accessed: 20.07.2017).
38. Bogolyubova O. N., Ledovaya. Y. A., Churilova A. G. Reprezentatsii psikhologicheskogo distressa v seti “Instagram” [Representations of psychological distress in Instagram]. *Anan’evskie chteniya — 2016: Psikhologiya: vchera, segodnya, zavtra: materialy mezhdunarodnoi nauchnoi konferentsii* [Ananyev readings — 2016: Psychology of yesterday, today, tomorrow: Proceedings of the international scientific conference]. Eds Shabolov A. V. et al. St. Petersburg, Aising Publ., 2016, vol. 2, pp. 125–126. (In Russian)
39. Padrez K. A. et al. Linking social media and medical record data: a study of adults presenting to an academic, urban emergency department. *BMJ quality & safety*, 2016, vol. 25, no. 6, pp. 414–423.
40. Penn Positive Psychology Center. *World well-being project*. Available at: <http://wwbp.org/about.html> (accessed: 22.08.2017).
41. John Templeton Foundation. *Grant Database*. 2017. Available at: <https://www.templeton.org/grants/grant-database> (accessed: 22.08.2017).
42. Gnip Inc. *Enterprise access to Twitter data*. Available at: <https://gnip.com/sources/twitter/> (accessed: 22.08.2017).
43. Preotiu-Pietro D., Carpenter J., Giorgi S., Ungar L. Studying the Dark Triad of personality through Twitter behavior. *Proceedings of the 25th ACM International Conference on Information and Knowledge Management — CIKM ’16, New York*. New York, ACM Press, 2016, pp. 761–770.
44. Park G. et al. Women are Warmer but No Less Assertive than Men: Gender and Language on Facebook. *PLOS ONE*, 2016, vol. 11, no. 5, p. e0155885.
45. Diener E., Emmons R. A., Larsen R. J., Griffin S. The Satisfaction With Life Scale. *Journal of Personality Assessment*, 1985, vol. 49, no. 1, pp. 71–75.
46. Schwartz H. A. et al. Predicting individual well-being through the language of social media. *Biocomputing 2016: Proceedings of the Pacific Symposium*, 2016, pp. 516–527.
47. Kosinski M., Stillwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences (PNAS)*, 2013, vol. 110, no. 15, pp. 5802–5805.

For citation: Ledovaya Y. A., Tikhonov R. V., Bogolyubova O. N. Social networks as a new environment for interdisciplinary studies of human behavior. *Vestnik SPbSU. Psychology and Education*, 2017, vol. 7, issue 3, pp. 193–210. <https://doi.org/10.21638/11701/spbu16.2017.301>

Статья поступила в редакцию 23 августа 2017 г.

Статья принята к публикации 7 сентября 2017 г.